

United States Patent and Trademark Office

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/992,865	11/19/2001 Larry Yen 2		20528/301	5994
7590 03/10/2006			EXAMINER	
Oppenheimer Wolff & Donnelly LLP			DESHPANDE, KALYAN K	
45 South Seventh Street, Suite 3300 Minneapolis, MN 55402-1609			ART UNIT	PAPER NUMBER
• •			3623	

DATE MAILED: 03/10/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

	Application No.	Applicant(s)				
	09/992,865	YEN ET AL.				
Office Action Summary	Examiner	Art Unit				
•	Kalyan K. Deshpande	3623				
The MAILING DATE of this communication app						
Period for Reply						
A SHORTENED STATUTORY PERIOD FOR REPLY WHICHEVER IS LONGER, FROM THE MAILING DA - Extensions of time may be available under the provisions of 37 CFR 1.13 after SIX (6) MONTHS from the mailing date of this communication. - If NO period for reply is specified above, the maximum statutory period w - Failure to reply within the set or extended period for reply will, by statute, Any reply received by the Office later than three months after the mailing earned patent term adjustment. See 37 CFR 1.704(b).	ATE OF THIS COMMUNICATION 36(a). In no event, however, may a reply be tim vill apply and will expire SIX (6) MONTHS from cause the application to become ABANDONE	l. lely filed the mailing date of this communication. D (35 U.S.C. § 133).				
Status						
1) Responsive to communication(s) filed on 19 No.	ovember 2001.					
,	This action is FINAL . 2b)⊠ This action is non-final.					
. —	Since this application is in condition for allowance except for formal matters, prosecution as to the merits is					
closed in accordance with the practice under Ex parte Quayle, 1935 C.D. 11, 453 O.G. 213.						
Disposition of Claims						
4) Claim(s) 1-25 is/are pending in the application. 4a) Of the above claim(s) is/are withdraw 5) Claim(s) is/are allowed. 6) Claim(s) 1-25 is/are rejected. 7) Claim(s) is/are objected to. 8) Claim(s) are subject to restriction and/or	vn from consideration.					
Application Papers						
9) The specification is objected to by the Examine 10) The drawing(s) filed on 19 November 2001 is/a Applicant may not request that any objection to the Replacement drawing sheet(s) including the correct 11) The oath or declaration is objected to by the Ex	re: a)⊠ accepted or b)⊡ object drawing(s) be held in abeyance. See ion is required if the drawing(s) is obj	e 37 CFR 1.85(a). ected to. See 37 CFR 1.121(d).				
Priority under 35 U.S.C. § 119						
 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f). a) All b) Some * c) None of: 1. Certified copies of the priority documents have been received. 2. Certified copies of the priority documents have been received in Application No. 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)). * See the attached detailed Office action for a list of the certified copies not received. 						
Attachment(s) 1) Notice of References Cited (PTO-892) 2) Notice of Draftsperson's Patent Drawing Review (PTO-948)	4)					
3) Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date		atent Application (PTO-152)				

DETAILED ACTION

Introduction

1. The following is a non-final office action in response to the communications received on November 19, 2001. Claims 1-25 are now pending in this application.

Claim Rejections - 35 USC § 112

- 2. The following is a quotation of the first paragraph of 35 U.S.C. 112:
 - The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.
- 3. Claims 1-25 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the enablement requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

Independent claims 1, 10, and 18 recite the limitation, "inspecting the subset of the worked data corresponding to the duplicated subset of the input data to determine the accuracy of the subset of worked data". Additionally, claims 2, 11, and 19 recite the limitation, "predict the quality of the worked data". Furthermore, claims 6, 15, and 23 recite the limitation, "identifying the subset of the worked data resulting from the duplicated subset of input data". These limitations are representative of subjective steps that may be performed in the mind of the user, thus raising the issue of abstract ideas that require undue experimentation for the invention to be performed. Since many of the steps of the claims use subjective questions to gather subjective answers, which

are evaluated subjectively and lack a concise formula or description for how to evaluate the answers, one skilled in the art would have to conduct undue experimentation in order to perform the invention. Therefore, claims 1-25 are considered as failing to comply with the enablement requirement.

Claim Rejections - 35 USC § 101

- 4. 35 U.S.C. 101 reads as follows:
 - Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.
- 5. Claims 1-25 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. The claimed invention is required to produce a useful, concrete, and tangible real-world result. An invention that fails to produce a tangible result is one that involves no more than the manipulation of an abstract idea. See *State Street Bank & Trust Co. v. Signature Financial Group Inc.*, 149 F. 3d 1368, 47 USPQ2d 1596 (Fed. Cir. 1998). In order to be concrete the result must be substantially repeatable or the process must substantially produce the same result again.

Claims 1, 10, and 18 merely recite the manipulation of an abstract idea and do not produce a concrete result. Claims 1, 10, and 18 recite "inspecting the subset of the worked data corresponding to the duplicated subset of the input data to determine the accuracy of the subset of worked data", which is a mere abstract idea that does not produce real-world results. The step of "inspecting the subset of the worked data corresponding to the duplicated subset of the input data to determine the accuracy of the subset of worked data" is based on subjective standards. The results of this step

will not produce concrete real-world results since there is no evidence that this step, when repeated, will produce substantially the same result. This step is based on a subjective standard and will produce different results for each individual performing the step. Furthermore, the results from the step of "inspecting the subset of the worked data corresponding to the duplicated subset of the input data to determine the accuracy of the subset of worked data" are not in a tangible form providing the user with a "real-world" result. The results from this mental step remain within the mind of the person performing the step. Because the results produced by the method are not tangible and concrete, claims 1, 10, and 18 are considered to be directed toward non-statutory subject matter.

Claims 2, 11, and 19 fail to remedy claims 1, 10, and 18 being directed towards non-statutory subject matter and further recite the manipulation of an abstract idea and do not produce a concrete result as well. Claims 2, 11, and 19 recite "predict the quality of the worked data", which is a mere abstract idea that does not produce real-world results. The step of "predict the quality of the worked data" is based on subjective standards. The results of this step will not produce concrete real-world results since there is no evidence that this step, when repeated, will produce substantially the same result. This step is based on a subjective standard and will produce different results for each individual performing the step. Furthermore, the results from the step of "predict the quality of the worked data" are not in a tangible form providing the user with a "real-world" result. The results from this mental step remain within the mind of the person performing the step. Because the results produced by the method are not tangible and

Application/Control Number: 09/992,865 Page 5

Art Unit: 3623

concrete, claims 2, 11, and 19 are considered to be directed toward non-statutory subject matter.

Claims 6, 15, and 23 fail to remedy claims 1, 10, and 18 being directed towards non-statutory subject matter and further recite the manipulation of an abstract idea and do not produce a concrete result as well. Claims 6, 15, and 23 recite "identifying the subset of the worked data resulting from the duplicated subset of input data", which is a mere abstract idea that does not produce real-world results. The step of "identifying the subset of the worked data resulting from the duplicated subset of input data" is based on subjective standards. The results of this step will not produce concrete real-world results since there is no evidence that this step, when repeated, will produce substantially the same result. This step is based on a subjective standard and will produce different results for each individual performing the step. Furthermore, the results from the step of "identifying the subset of the worked data resulting from the duplicated subset of input data" are not in a tangible form. The results from this mental step remain within the mind of the person performing the step. Because the results produced by the method are not tangible and concrete, claims 6, 15, and 23 are considered to be directed toward non-statutory subject matter.

Claims 3-5, 7-9, 12-14, 16-17, 20-22, and 24-25 recite the same non-statutory subject matter as claims 1, 10 and 18 and fail to remedy these claims from being directed towards non-statutory subject matter and therefore are rejected as well.

Claim Rejections - 35 USC § 103

Application/Control Number: 09/992,865 Page 6

Art Unit: 3623

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

- (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.
- 7. Claims 1-25 are rejected under 35 U.S.C. 103(a) as being unpatentable over Khandekar (U.S. Patent No 6732102) further in view of Aragon (U.S. Patent No. 6055327).

As per claim 1, Khandekar teaches:

A method for testing output quality from a data extraction process, comprising: receiving input data containing information to be inserted into a database (see column 19 lines 4-43; where a designer selects a data source and data from this data source is retrieved and received by the system.);

dividing the input data into a plurality of batches such that a subset of the input data is duplicated among the plurality of batches (see column 11 lines 51-60 and column 18 lines 49-65; where selected data to be captured is copied and saved in to internal files.);

receiving the worked data from each of the plurality of data entry clerks (see column 18 lines 49-65 and column 20 lines 50-67; where the worked data is completed and available to end users.); and

inspecting the subset of the worked data corresponding to the duplicated subset of the input data to determine the accuracy of the subset of worked data (see

Art Unit: 3623

column 18 lines 49-65 and column 20 lines 50-67; where the worked data can be inspected.);.

Khandekar fails to explicitly teach "distributing the plurality of batches to a plurality of data entry clerks, wherein each data entry clerk processes one of the plurality of batches and converts data from the batch into worked data". Khandekar does teach users determining data to be processed and setting the system to capture the data and format it in to working data (see column 19 lines 4-43). In other words, Khandekar teaches the automation of the manual process of this invention. The advantages of making the process of processing data in to working data is that it allows for greater accuracy based on human monitoring. It would have been obvious, at the time of the invention, to make manual the processing of data to the automated feature of Khandekar in order to increase accuracy due to human monitoring, which is a goal of Khandekar (see column 2 lines 15-42). Furthermore, the Courts have held that the automation of a process is within ordinary skill level in the art. See *In re Venner*, 120 USPQ 192, 194; 262 F2d 91 (CCPA 1958). The making manual of an automated process will also be within the ordinary skill in the art.

As per claim 2, Khandekar fails to teach "the step of inspecting predicts the quality of the worked data". Aragon teaches "the step of inspecting predicts the quality of the worked data" (see column 15 lines 8-41; where the operator decides whether the sample of records inspected verifies that the batch does not require additional inspection or whether the operator requires additional samples to make this determination.). The advantage of predicting the quality of work is that it increases the

Art Unit: 3623

accuracy of the compiled data. It would have been obvious, at the time of the invention, for one of ordinary skill in the art to combine the feature of predicting the quality of the worked data from the Aragon system to the Khandekar system in order to increase the accuracy of the worked data, which is a goal of Khandekar (see column 2 lines 15-42).

As per claim 3, Khandekar teaches:

The method for testing output quality from claim 1, wherein the subset of the input data duplicated among the batches is based on a sampling plan (see column 20 lines 11-35; where the input data is gathered based on a schema determined by the user.).

As per claim 4, Khandekar teaches:

The method for testing output quality from claim 1, further comprising repeating the steps of dividing, distributing, receiving and inspecting, if a desired level of accuracy is not reached (see column 21 lines 35-53; where the program can be repeated at the user's discretion.).

As per claim 5, Khandekar fails to teach "adjusting the desired level of accuracy based on inspecting the subset of the worked data". Aragon teaches "adjusting the desired level of accuracy based on inspecting the subset of the worked data" (see column 15 lines 42-67 and column 16 lines 1-5; where the accuracy of the data can be adjusted by different methods including enforcing greater accuracy of the operator or using a more experienced operator.). The advantage of adjusting the desired level of accuracy is that it allows for the user to make the system more efficient. It would have been obvious, at the time of the invention, for one of ordinary skill in the art to combine

Art Unit: 3623

the feature of adjusting the level of accuracy from the Aragon system to the Khandekar system in order to make the system more efficient, which is a goal of Khandekar (see column 2 lines 31-42).

As per claim 6, Khandekar fails to teach "the step of inspecting the subset of the worked data comprises: identifying the subset of the worked data resulting from the duplicated subset of the input data; comparing entries made by each of the plurality of data clerks on the subset of the worked data; and flagging the entries that differ. Aragon teaches "the step of inspecting the subset of the worked data comprises: identifying the subset of the worked data resulting from the duplicated subset of the input data; comparing entries made by each of the plurality of data clerks on the subset of the worked data; and flagging the entries that differ" (see column 13 lines 58-67 and column 14 lines 1-30; where the operator uses a sample of data records, compares the data values, and marks erroneous records.). The advantage of identifying and flagging differing data entries is that it increases the accuracy of the processed data. It would have been obvious, at the time of the invention, for one of ordinary skill in the art to combine the feature of identifying and flagging differing data entries from the Aragon system to the Khandekar system in order to increase the accuracy of the processed data, which is a goal of Khandekar (see column 2 lines 15-42).

As per claim 7, Khandekar fails to teach "the step of inspecting the subset of the worked data comprises: accepting the worked data for submission to a database if the desired level of accuracy is met and rejecting the worked data for submission to the database if the desired level of accuracy is not met". Aragon teaches "the step of

inspecting the subset of the worked data comprises: accepting the worked data for submission to a database if the desired level of accuracy is met and rejecting the worked data for submission to the database if the desired level of accuracy is not met" (see column 15 lines 8-41; where the operator determines whether a batch of records is to be accepted.). The advantage of accepting accurate information and rejecting erroneous information is that it increases the accuracy of the information stored. It would have been obvious, at the time of the invention, for one of ordinary skill in the art to combine the feature of accepting accurate information and rejecting erroneous information from the Aragon system to the Khandekar system in order to increase the accuracy of the stored information, which is a goal of Khandekar (see column 2 lines 15-42).

As per claim 8, Khandekar teaches:

The method for testing output quality from claim 1, wherein the input data is a plurality of technical product data sheets (see column 8 lines 30-39; where the input data is technical product data on a web page. The example provided is technical information on stock accounts on web pages, which are data sheets.).

As per claim 9, Khandekar teaches:

The method for testing output quality from claim 1, wherein the steps of dividing, distributing, receiving and inspecting are accomplished with a computer system (see column 7 lines 10-62; where the data extraction system runs on a computer system.).

As per claim 10, Khandekar teaches:

Art Unit: 3623

A data extraction tool implemented on a computer, the tool comprising:

a first receiver unit for receiving input data containing information to be inserted into a database (see column 19 lines 4-43; where a designer selects a data source and data from this data source is retrieved and received by the system.);

a data divider unit for dividing the input data into a plurality of batches such that a subset of the input data is duplicated among the plurality of batches (see column 11 lines 51-60 and column 18 lines 49-65; where selected data to be captured is copied and saved in to internal files.);

a distributor unit for distributing the plurality of batches to a plurality of data entry clerks, wherein each data entry clerk processes one of the plurality of batches and converts data from the batch into worked data;

a second receiver unit for receiving the worked data from each of the plurality of data entry clerks (see column 18 lines 49-65 and column 20 lines 50-67; where the worked data is completed and available to end users.); and

an inspector unit for inspecting the subset of the worked data corresponding to the duplicated subset of the input data to determine the accuracy of the subset of worked data (see column 18 lines 49-65 and column 20 lines 50-67; where the worked data can be inspected.).

Khandekar fails to teach "a distributor unite for distributing the plurality of batches to a plurality of data entry clerks, wherein each data entry clerk processes one of the plurality of batches and converts data from the back in to worked data. This limitation is

Art Unit: 3623

already addressed by the rejection of claim 1; therefore the same rejection applies to this claim.

As per claim 11, Khandekar fails to teach wherein the inspector unit predicts the quality of the worked data. This limitation is already addressed by the rejection of claim 2; therefore the same rejection applies to this claim.

As per claim 12, Khandekar teaches:

The data extraction tool implemented on a computer from claim 10, wherein the subset of the input data duplicated among the batches is based on a sampling plan (see column 20 lines 11-35; where the input data is gathered based on a schema determined by the user.).

As per claim 13, Khandekar teaches:

The data extraction tool implemented on a computer from claim 10, further comprising reworking the batch using the distributor unit, second receiver unit, and inspector unit, if a desired level of accuracy is not reached (see column 21 lines 35-53; where the program can be repeated at the user's discretion.).

As per claim 14, Khandekar fails to teach "adjusting the desired level of accuracy based on the inspector unit inspecting the subset of the worked data". This limitation is already addressed by the rejection of claim 5; therefore the same rejection applies to this claim.

As per claim 15, Khandekar fails to teach "the inspecting of the subset of the worked data performed by the inspector unit comprises: identifying the subset of the worked data resulting from the duplicated subset of the input data; comparing entries

Art Unit: 3623

made by each of the plurality of data clerks on the subset of the worked data; and flagging the entries that differ". This limitation is already addressed by the rejection of claim 6; therefore the same rejection applies to this claim.

As per claim 16, Khandekar fails to teach "the inspecting of the subset of the worked data performed by the inspector unit comprises: accepting the worked data for submission to a database if the desired level of accuracy is met and rejecting the worked data for submission to the database if the desired level of accuracy is not met". This limitation is already addressed by the rejection of claim 7; therefore the same rejection applies to this claim.

As per claim 17, Khandekar teaches:

The data extraction tool implemented on a computer from claim 10, wherein the input data is a plurality of technical product data sheets (see column 8 lines 30-39; where the input data is technical product data on a web page. The example provided is technical information on stock accounts on web pages, which are data sheets.).

As per claim 18, Khandekar teaches:

A computer program for a data extraction tool, the computer program embodied on a computer readable medium for execution by a computer, the computer program comprising:

a code segment that receives input data containing information to be inserted into a database (see column 19 lines 4-43; where a designer selects a data source and data from this data source is retrieved and received by the system.);

a code segment that divides the input data into a plurality of batches such that a subset of the input data is duplicated among the plurality of batches (see column 11 lines 51-60 and column 18 lines 49-65; where selected data to be captured is copied and saved in to internal files.);

a code segment that receives the worked data from each of the plurality of data entry clerks (see column 18 lines 49-65 and column 20 lines 50-67; where the worked data is completed and available to end users.); and

a code segment that inspects the subset of the worked data corresponding to the duplicated subset of the input data to determine the accuracy of the subset of worked data (see column 18 lines 49-65 and column 20 lines 50-67; where the worked data can be inspected. Per the Specification page 10, the inspection software assists the inspector by displaying information side by side.).

Khandekar fails to explicitly teach "a code segment that distributes the plurality of batches to a plurality of data entry clerks, wherein each data entry clerk processes one of the plurality of batches and converts data from the batch into worked data". This limitation is already addressed by the rejection of claim 1; therefore the same rejection applies to this claim.

As per claim 19, Khandekar fails to teach "the code segment that inspects the data predicts the quality of the worked data". This limitation is already addressed by the rejection of claim 2; therefore the same rejection applies to this claim.

As per claim 20, Khandekar teaches:

The computer program for a data extraction tool from claim 18, wherein the subset of the input data duplicated among the batches is based on a sampling plan (see column 20 lines 11-35; where the input data is gathered based on a schema determined by the user.).

As per claim 21, Khandekar teaches:

The computer program for a data extraction tool from claim 18, further comprising reworking the batch using the code segment that distributes, the code segment that receives, and the code segment that inspects, if a desired level of accuracy is not reached (see column 21 lines 35-53; where the program can be repeated at the user's discretion.).

As per claim 22, Khandekar fails to teach "adjusting the desired level of accuracy based the code segment that inspects inspecting the subset of the worked data". This limitation is already addressed by the rejection of claim 5; therefore the same rejection applies to this claim.

As per claim 23, Khandekar fails to teach "the step of inspecting performed by the code segment that inspects comprises: identifying the subset of the worked data resulting from the duplicated subset of the input data; comparing entries made by each of the plurality of data clerks on the subset of the worked data; and flagging the entries that differ". This limitation is already addressed by the rejection of claim 6; therefore the same rejection applies to this claim.

As per claim 24, Khandekar fails to teach "the step of inspecting performed by the code segment that inspects comprises: accepting the worked data for submission to

Art Unit: 3623

a database if the desired level of accuracy is met and rejecting the worked data for submission to the database if the desired level of accuracy is not met". This limitation is already addressed by the rejection of claim 7; therefore the same rejection applies to this claim.

As per claim 25, Khandekar teaches:

The computer program for a data extraction tool from claim 18, wherein the input data is a plurality of technical product data sheets (see column 8 lines 30-39; where the input data is technical product data on a web page. The example provided is technical information on stock accounts on web pages, which are data sheets.).

Conclusion

8. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. The following are pertinent to the current invention, though not relied upon:

De Pauw et al. (U.S. Patent No. 6370684) teaches methods for extracting reference patterns in JAVA and depicting the same.

Aptroot-Soloway (U.S. Patent No. 3974496) teaches a method for the presentation of at least two different sets of information to an observer simultaneously and in the same location, one superposed upon the other, one set of information concerning characters inscribed on a source carrier or document, the other set produced by a data processing machine

Woo et al. (U.S. Patent No. RE35738) teaches a data entry and error embedding system in which, first, a document is bitmapped and recorded in a first memory.

Zabih et al. (U.S. Patent No. 6181817) teaches a method of comparing data objects using joint histograms.

Graham et al. (U.S. Patent No. 6411974) teaches a method extracts desired contents from multiple heterogeneous textual streams and provides normalized data representative of the desired contents.

Liddle et al. (Liddle, Stephen W.; Campbell, Douglas M.; Crawford, Chad; "Automatically Extracting Structure and Data from Business Reports", INT CONF INF KNOWLEDGE MANAGE., 1999) teaches data mining of business reports and algorithms for the extraction of information from business reports.

Embley et al. (Embley, D.W.; Campbell, D.M.; Jiang, Y.S.; Liddle, S.W.; Lonsdale, D.W.; Ng, Y.K.; Smith, R.D.; "Conceptual-Model-Based Data Extraction from Multiple-Record Web Pages", *Data and Knowledge Engineering*, 1999) teach a conceptual based data extraction method for extracting data from web pages.

Taghva et al. (Taghva, Kazem; Borsack, Julie; Condit, Allen; "Evaluation of Model-Based Retrieval Effectiveness with OCR Text", *ACM Transactions on Information Systems*, January, 1996, pp. 64-93) teaches the accuracy of data retrieval using OCR.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Kalyan K. Deshpande whose telephone number is (571) 272-5880. The examiner can normally be reached on M-F 8am-5pm.

Art Unit: 3623

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Tariq Hafiz can be reached on (571) 272-6729. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Kasa Ry

TARIQ R. NAFIZ SUPERVISORY PATENT EXAMINER

LECHWOFOUR CENTRES :